Method of Genotyping by Hybridisation Analysis

This invention relates to methods of genotyping individuals, in particular to methods for assigning a genotype to a nucleic acid sample obtained from an individual by analysis of the hybridization of the nucleic acid sample with a nucleic acid probe.

Genotyping methods commonly involve the production of melting or annealing curves for the duplexes formed when a sample binds to a nucleic acid probe. Curves of the primary or derivative melting/annealing data are then analysed in order to classify the tested sample into a genotype category.

Various algorithms are commonly used to carry out this analysis, using either limited information from each curve (e.g., comparison of peak positions) or substantial information from each curve (e.g., comparison of polynomial or other mathematical models for the curves).
However, these techniques are slow and have a high error rate.

The present inventors have identified a rapid and simple method of genotyping nucleic acid samples.

One aspect of the invention provides a method of genotyping a nucleic acid sample comprising;

(a) providing melt curves for a population of nucleic acid samples hybridized to one or more nucleic acid probes;

(b) applying one or more grouping lines to the population of melt curves,

wherein each of the grouping lines intersects a number of melt curves within said population,

(c) assigning genotype categories to said grouping lines, and;

(d) determining the genotype category of a nucleic acid sample in said population by identifying one or more grouping lines which intersect the melt curve of the sample.

The method may include an initial step of obtaining the melt curve data from a population of nucleic acid samples. Melt curves may be obtained for a population of nucleic acid samples by;

(a) contacting a population of nucleic acid samples with one or more nucleic acid probes which hybridize with each of the samples to form a population of complexes,

(b) progressively altering the hybridization conditions to decrease or increase the formation of said complexes;

(c) measuring output signals indicative of the extent of hybridization of the complexes; and,

(d) plotting the output signals relative to the hybridization conditions for each of said population of complexes to produce a population of melt curves.

A melt curve may be a melting curve obtained by increasing the stringency of the hybridization conditions and monitoring the dissociation of the sample and the probe nucleic acids or an annealing curve obtained by reducing the stringency of the conditions and monitoring the association of the sample and the probe nucleic acids.

The output signal is indicative of the degree or amount of formed duplex and may either increase or decrease as the amount of formed duplex increases in an annealing reaction or decreases in a dissociation reaction. Many suitable signal mechanisms are known in the art.

The stringency of the hybridization conditions may be progressively altered by increasing or decreasing a parameter of hybridization stringency, such as temperature, voltage or pH.

Methods and procedures for obtaining melt curves for nucleic acid complexes are well-known in the art (see, for example US6,174,670, US5,789,167, Drobyshev et al Gene 188 (1997) 45 52, Kochinsky and Mirzabekov Human Mutation (2002) 19:343-360, Livshits et al J. Biomol. Structure Dynam. (1994) 11 783-795, Howell et al (1999) Nature Biotechnology 17 87-88.

The nucleic acid sample is preferably a single stranded nucleic acid having at least 30, at least 40 or at least 50 nucleotides. In some preferred embodiments, the nucleic acid sample may have less than 250 nucleotides, less than 200 nucleotides or less than 150 nucleotides.

Nucleic acid for use in the present methods may be wholly or partially synthetic and may include genomic DNA, cDNA, RNA and analogues thereof, such as PNA and LNA (locked nucleic acid). A nucleic acid may also comprise one or more labelled or modified bases. Modified bases may include, for example, acetylcytidine, methylcytidine, dihydrouridine, methylpseudouridine, D-galactosylqueosine, methyladenosine, methylpseudouridine, methylguanosine, methylinosine, D-mannosylqueosine, wybutoxosine, pseudouridine, queosine, thiocytidine and thiouridine.

A nucleic acid strand may also comprise one or more modifiers of base-pair stability.

The nucleic acid sample may be obtained by amplifying a region of sample DNA containing one or more positions or sites of variation. A single strand of the amplified product may then be isolated and/or purified.

A position of variation is a position within the tested sequence which may differ between samples obtained from different individuals or between cells within an individual, for example due to allelic variation, polymorphism or mutation. For example, there may be an insertion, deletion or substitution of one or more nucleotides at a position of variation, relative to an allelic reference sequence. In some embodiments, the sample may contain a polymorphism at a position of variation, for example a single nucleotide polymorphism.

Sample nucleic acid from an individual may be subjected to a specific amplification reaction such as the polymerase chain reaction (PCR) (reviewed for instance in "PCR protocols; A Guide to Methods and Applications", Eds. Innis et al, 1990, Academic Press, New York, Mullis et al, Cold Spring Harbor Symp. Quant. Biol., 51:263, (1987), Ehrlich (ed), PCR technology, Stockton Press, NY, 1989, and Ehrlich et al, Science, 252:1643-1650, (1991)) to generate the nucleic acid sample. DNA amplification using techniques such as PCR are well-known in the art.

The nucleic acid probe may comprise or consist of the nucleotide sequence of the most common allele of the position of variation or one of the alternative allele sequences, as a reference. Suitable nucleic acid probes may consist of at least 10 nucleotides, at least 13 nucleotides, at least 15 nucleotides, at least 20 nucleotides or at least 25

nucleotides. The probe may be designed such that the variant position is anywhere within the probe sequence. Preferably, the variant position is located within the central third of the probe sequence, more preferably at the central position within the probe sequence.

The nucleic acid sample and/or the nucleic acid probe may be bound to a 2-dimensional solid surface, placed or immobilised within a 3-dimensional matrix, or free in solution in accordance with the particular method employed to obtain the melt curve.

Various techniques for synthesizing oligonucleotide probes are well known in the art, including phosphotriester and phosphodiester synthesis methods.

In some embodiments, a detectable label may be covalently or non-covalently attached to the probe and/or the sample using conventional techniques. The label preferably mediates the production of an output signal which is indicative of the extent of hybridization of the nucleic acid complex. The output signal may, for example, be produced or modulated i.e. increased or decreased, in the presence of the hybridization complex relative to its absence, or in the absence of the hybridization complex relative to its presence.

Suitable output signals include indicators such as luminescent (fluorescence/phosphorescence) intensity or decay time measurements, light polarization measurements, light absorption/transmission and reflectance measurements, chemiluminescence signals, scattered light patterns and evanescent fields. Methods for the production and measurement of these output signals are well-known in the art.

A melt curve for a hybridization complex may be obtained by any convenient method. Many such methods are known in the art.

A melt curve may consist of a graphic plot or display of the variation of the output signal with the parameter of hybridization stringency. Output signal may be plotted directly against the hybridization parameter. Typically, a melt curve will have the output signal, for example fluorescence, which indicates the degree of duplex structure (i.e. the extent of hybridization), plotted on the Y-axis and the hybridization parameter on the X axis.
In other embodiments, melt curves may be provided by plotting the first or the second derivative of the output signal (or the negative values thereof) against the hybridization parameter.

Preferred hybridization parameters include temperature (or time, if temperature was altered steadily), pH and voltage.

In preferred embodiments, the population of melt curves is normalised to a common start point prior to applying said grouping lines.

The population of melt curves may comprise one or more reference melt curves. Reference melt curves may be obtained from the interaction of a nucleic acid probe with a sample sequence of a known genotype.

A grouping line may be any one-dimensional shape which may be applied in a graphic form to the population of curves. In preferred embodiments, a grouping line is a one-dimensional

shape i.e. a straight or curved line, preferably a straight line.

Grouping lines may be applied to the population of melt curves by determining the presence of a discrete cluster of melt curves within the population which are distinct from other melt curves within the population. A discrete cluster may comprise a plurality of curves which are positioned closer to each other for all or a discrete range of X and/or Y axis values, than to other curves in the population. Preferably, each such cluster comprises a reference curve. A grouping line is applied so as to intersect only the melt curves within the cluster and may intersect any part or region of a melt curve.

In some embodiments, grouping lines may be applied to clusters which comprise a reference curve. In other embodiments, grouping lines may be applied to a population of melt curves in accordance with results obtained from the analysis of previous populations of melt curves, for example the known positions of reference melt curves within those previous populations. Grouping lines which represent particular genotype categories may thus be applied to a displayed population of curves at the same positions as on previously displayed curve populations produced using the same experimental system. In other words, grouping lines may be used repetitively for a particular system once they have been established for that system. There is therefore no requirement for each melt curve population to comprise reference melt curves.

The grouping lines may be applied to the population of melt curves by an operator in order to categorise the samples. In

preferred embodiments, the population of melt curves is provided by a data processing means and displayed on a monitor or other image display. The operator may then apply the grouping lines to the displayed images manually by means of a graphic interface. For example, the operator may apply grouping lines to the melt curves displayed on the monitor using a keypad, mouse, touchpad, trackball, pressure-sensitive stylus, or other interface device. Suitable graphic interfaces and interface devices are well-known in the art.

In other embodiments, the grouping lines may be applied to population of melt curves automatically by a data processor. Many different strategies for applying the grouping lines are possible and can be readily implemented by those skilled in the art. For example, the data processor may be adapted to apply grouping lines by;

(i) tracking the Y-value distribution of said melt curves along the X-axis, (ii) identifying one or more regions in which said melt curves separate into distinct clusters, and (iii) applying one or more grouping lines to define each said cluster.

Alternatively, the data processor may be adapted to apply grouping lines by; (i) applying a plurality of candidate lines to the population of melt curves, and (ii) identifying one or more candidate lines which only intersect a discrete cluster of curves within said population as grouping lines.

In other embodiments, grouping lines may be applied by;

(i) retrieving the stored positions of one or more established grouping lines;

(ii) applying said established grouping lines to a displayed population of curves.

In general, a nucleic acid sample from a diploid genome may have three possible genotypes at a position of variation. The sample may be homozygous for a match with a reference allelic sequence, homozygous for a mismatch with a reference allelic sequence, or heterozygous. Grouping lines may thus be assigned to the following genotypes; homozygotes for matched sequences, homozygotes for mismatched sequences, and heterozygotes. Further genotype categories are also possible if the sample contains more than one position of variation or is present in multiple copies in the genome.

A nucleic acid sample from a haploid genome may have two possible genotypes at a position of variation. The sample may match a reference allelic sequence (i.e. 'homozygous' for a match with the reference sequence) or mismatch a reference allelic sequence (i.e. 'homozygous' for a mismatch with the reference sequence). Further genotype categories are also possible if the sample contains more than one position of variation or is present in multiple copies in the genome.

A nucleic acid sample from a polyploid genome may have more than three genotypes at a position of variation. In addition to being homozygous for a match or a mismatch with a reference allelic sequence, the sample may have several heterozygous genotypes, depending on the number of match and mismatch alleles present in the genome. Further genotype categories are also possible if the sample contains more than one position of variation or is present in multiple copies in the genome.

As described above, each grouping line or combination of grouping lines may be assigned to a genotype category i.e. grouping lines are applied to the distinct clusters in the population of curves and then assigned to each genotype category. In other embodiments, grouping lines preassigned to each genotype category may be applied to distinct melt curve clusters in the population of curves. For example, a grouping line may be applied to intersect a cluster of curves that comprises a reference curve of a known genotype.

Grouping lines may be assigned to a genotype category by determining the context of the melt curves intersected by the line within the population of curves. For example, a grouping line which intersects melt curves that show only a high temperature of melting relative to the population as a whole may be assigned to samples which are homozygotes for sequences that matched the probe. A grouping line which intersects melt curves that show only a low temperature of melting relative to the population as a whole may be assigned to samples which are homozygotes for sequences that mismatched the probe. A grouping line which intersects melt curves that show both a high and a low temperature of melting may be assigned to samples which are heterozygotes.

In some embodiments, a grouping line may be assigned to the genotype category of a reference curve that is intersected by the grouping line. In other embodiments, a grouping line may be applied or assigned on the basis of its position relative to the known reference curves from previous experiments using the same system.

In some circumstances, for example, with assays or platforms which give rise to less than ideal melt curves, target

sequences that do not react as single copy sequences, or assays with multiple probes, the population of melt curves may not contain distinct melt curve clusters which separate uniquely at any one position on the graphic display.

The clusters may be placed into genotype categories by determining the relative separation of the curves within the clusters at more than one region of the display, for example by applying grouping lines at each of these regions.

The assignment of the genotype category may be made with reference to one or more reference curves of known genotype. For example, the intersection of grouping lines with reference melt curves may provide a system or algorithm for the assignment of genotype category to a sample melt curve, based on the intersection of that curve with one or more grouping lines. Any algorithm or system is dependent on particular assay, platform, probe combination, and the target sequence in question.

A method may comprise;

(a) providing reference melt curves for a population of nucleic acid molecules of one or more known genotype categories hybridized to a nucleic acid probe;

(b) applying one or more grouping lines to the population of reference melt curves,

wherein each of the grouping lines intersects a one or more reference melt curves within said population,

(c) determining the intersection of the grouping lines by reference melt curves of each of the genotype categories; and,

d) providing an assignment algorithm which relates each genotype category to the intersected grouping lines.

The genotype category of a nucleic acid sample in a population may then be determined by applying the one or more grouping lines previously applied to the reference curves as described above to a population of sample melt curves, identifying which of the one or more grouping lines intersect the melt curve of the nucleic acid sample, and applying said assignment algorithm.

For example, a sample melt curve in an assay may intersect two or more grouping lines. The curve may be assigned to a genotype category by the implementation of an algorithm that relates the grouping lines intersected by the curve to a genotype category. Use of such an algorithm obviates the need for reference curves in every population of melt curves.

In some embodiments, an assignment system or algorithm may consist of allocating an order of precedence to the grouping lines. The sample curve is then genotyped according to the genotype category of the grouping line with the highest precedence. A method may comprise;

applying a plurality of grouping lines to the population of melt curves, and;

assigning an order of precedence to the plurality of grouping lines,

wherein samples having a melt curve which intersects two or more grouping line are assigned to the genotype category of the grouping line with the highest precedence.

In addition to the analysis of melt curves, methods of the invention may find more general application in the analysis and categorization of datasets.

Another aspect of the invention provides a method of categorising one or more datasets within a population of datasets comprising;

(a) producing a graphic plot for each of the datasets in said population to produce a population of graphic plots,

(b) applying one or more grouping lines to the population of graphic plots,

wherein each of the grouping lines intersects one or more plots within said population,

(c) assigning said grouping lines to categories, and;

(d) determining the category of one or more of the datasets in said population by determining the grouping line which intersects the graphic plot of the one or more datasets.

Preferably, a dataset consists of a series of values of an output signal indicative of the association or dissociation of two or more components of a biological complex in response to a changing physical parameter, such as temperature. Such a dataset may be plotted as a melt curve.

Melt curves may be obtained for a population of biological complexes by;

(a) contacting, in a medium, biological components which bind together to form a population of biological complexes,

(b) progressively altering the conditions of the medium to decrease or increase the formation of said complexes;

(c) measuring output signals indicative of the extent of formation of the complexes; and,

(d) plotting the output signals relative to the conditions for each of said population of complexes to produce a population of melt curves.

Suitable biological components include biological molecules such as polypeptides, nucleic acids, lipids and carbohydrates and biological particles such as microbial or eukaryotic cells and viral particles.

A biological complex may, for example, include an antibody/antigen complex, a multi-component protein or protein complex, a receptor/ligand complex, an enzyme/inhibitor complex or a double stranded nucleic acid complex.

Suitable conditions that may be altered include pH, temperature and electrical field strength.

A method of categorising one or more biological complexes or components thereof within a population may comprise;

(a) providing a population of melt curves of samples of biological complexes,

(b) applying one or more grouping lines to the population of melt curves,

wherein each of the grouping lines intersects one or more plots within said population,

(c) assigning said grouping lines to categories, and;

(d) determining the category of one or more of the samples in said by determining the grouping line which intersects the melt curve of the one or more samples.

The provision of melt curves and the application and assignment of grouping lines is described above, with reference to nucleic acid genotyping.

A biological complex or component thereof may be categorised in accordance with any parameter which affects complex association/dissociation. For example, a category may be an

allelic group (e.g. for polypeptides and complexes thereof), a species, sub-type or strain (e.g. for a viral particle or a microbial cell), an immunogenic class (e.g. for an antigen or antibody), or a genotype (e.g. for a nucleic acid).

Further aspects of the invention provide: (i) computer-readable code for performing a method described herein, (ii) a computer program product carrying such computer-readable code, and (iii) a computer system configured to perform a method described herein.

The term "computer program product" includes any computer readable medium or media which can be read and accessed directly by a computer. Typical media include, but are not limited to: magnetic storage media such as floppy discs, hard disc storage medium and magnetic tape; optical storage media such as optical discs or CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media.

A typical computer system of the present invention comprises a central processing unit (CPU), input means, output means and data storage means (such as RAM). A monitor or other image display is preferably provided.
For example, a computer system may comprise a processor adapted to perform a method of the invention. For example the processor may be adapted;

(a) to produce a graphic plot for each of the datasets in a population to produce a population of graphic plots,

(b) to display said population of plots

(c) to apply or to allow the user to apply one or more grouping lines to the displayed population of graphic plots,

wherein each of the grouping lines intersects one or more plots within said population,

(d) to assign each said grouping line to a category, and;

(e) to determine the category of one or more of the datasets in said population by determining the grouping line which intersects the graphic plot of the one or more datasets.

The datasets may be entered into the processor via the input means.

As described above, a dataset may consist of a series of values for an output signal which is indicative of the association or dissociation of two or more components of a biological complex in response to a progressively changing physical parameter, such as temperature. Such a dataset may be plotted as a melt curve. Biological complexes are described in more detail above.

In particular, a computer system according to the invention may comprise a processor adapted;

(a)to produce melt curves from annealing/denaturing data from a population of nucleic acid samples hybridized to one or more nucleic acid probes;

(b) to display said population of melt curves

(c) to apply or to allow the user to apply one or more grouping lines to the displayed population of melt curves,

wherein each of the grouping lines intersects a cluster of melt curves within said population,

(d) to assign each said grouping line to a genotype category, and;

(e) to determine the genotype category of a nucleic acid sample in said population by identifying the grouping line which intersects the melt curve of the sample.

In some embodiments, the user may apply the grouping lines to the displayed population of melt curves via the input means.

In other embodiments, the processor may adapted to apply grouping lines by

(i) tracking the Y-value distribution of said plots along the X-axis,

(ii) identifying one or more regions in which said plots separate into distinct clusters; and,

(iii) applying one or more grouping lines to define each said cluster.

In other embodiments, the processor may be adapted to apply grouping lines by;

(i) applying a plurality of candidate lines to said population of plots, and;

(ii) identifying one or more candidate lines which only intersect a discrete cluster of curves within said population as grouping lines.

In other embodiments, the processor may store the positions of grouping lines which have been assigned to genotype categories on the basis of a population of reference curves and apply grouping lines at the stored positions on a population of sample curves. Thus, the processor may be adapted to apply grouping lines by;

(i) retrieving the stored positions of one or more established grouping lines;
(ii) applying said established grouping lines to a displayed population of curves.

Annealing/denaturing data from a population of nucleic acid samples hybridized to a nucleic acid probe may include the amount of an output signal at one or more different proteins and protein complexes, nucleic acid materials, acceptors and receptors, peptides, lectins, saccharides, carbohydrates, lipids and lipid complexes, other biological macromolecules and complexes, ligands such as antigens, haptens, inhibitors, agonists, and antagonists, and even membranes, organelles, cells, and tissues, hybridization conditions for a specific nucleic acid sample, wherein the output signal being indicative of the extent of hybridization of a complex comprising a nucleic acid sample hybridized to a nucleic acid probe. Suitable output signals are described above.

The computer system may further comprise a memory device for storing the annealing/denaturing data. The genotype information may be stored on another or the same memory device, and/or may be sent to an output device or displayed on a monitor. The memory device may also store the positions of established grouping lines for a particular experimental system (i.e. a particular probe and assay format).

Another aspect of the invention provides a DNA hybridization device having an output signal detector and a computer system as described above for analyzing data obtained by the detector.

The output signal detector may, for example, detect luminescent (fluorescence/phosphorescence) intensity or decay time measurements, light polarization measurements, light absorption/transmission and reflectance measurements, chemiluminescence signals, scattered light patterns or evanescent fields. Suitable detectors are well known in the art. Techniques for measuring these output signals are routine in the art.

The DNA hybridization device may further comprise a hybridization chamber suitable for annealing or denaturing nucleic acid complexes in accordance with the invention. The device may comprise means for progressively altering the hybridization conditions within the chamber, for example by altering the temperature, pH or voltage.

Various further aspects and embodiments of the present invention will be apparent to those skilled in the art in view of the present disclosure. All documents mentioned in this specification are incorporated herein by reference in their entirety.

Certain aspects and embodiments of the invention will now be illustrated by way of example and with reference to the figures described below.

Figure 1 shows DNA melt curves of primary non-normalised data with applied grouping lines. Solid lines show homozygous matches to the probe, thick-hatched lines show homozygous mismatches and thin-hatched lines show heterozygotes.

Figure 2 shows DNA melt curves of primary normalised data with applied grouping lines.

Figure 3 shows DNA melt curves of negative first derivative non-normalised data with applied grouping lines.

Figure 4 shows DNA melt curves of negative first derivative normalised data with applied grouping lines.

Experimental

Materials and Methods

DNA Samples

Genomic DNA was prepared from blood samples of twelve unrelated Swedish females by standard DNA extraction protocols using organic solvents.

Target DNA Sequence

A known human single nucleotide polymorphism (SNP) marker was examined. This SNP exists in the gene *CLECSF6*, and is represented in the dbSNP database (Nucleic Acids Research, 2001, 29:308-311) under the unique identifier rs1133104. The 5′-3′ sequence of marker rs1133104 is as follows: CCGGATAAAAATTAAGAGAGACTCAtgttctctcc(a/c)tgtgtGCACAATGGGGGCCTCC AGTAC. Capitalized bases represent sequences used to prime Polymerase Chain Reaction (PCR) amplification, and the bracketed sequences represent the alternative alleles that exist at the polymorphic position.

Polymerase Chain Reaction

Employed PCR primers were rs1133104b01A (5'-Biotinylated oligonucleotide 5′-GTACTGGAGGCCCCCATTGTGC-3′) and rs1133104-01B (oligonucleotide 5′-CCGGATAAAAATTAAGAGAGACTCA-3′). Reactions were in 5 µl volume, using 1ng of genomic DNA, 0.75 pmol rs1133104b01A, 3 pmol rs1133104-01B, 0.03 units AmpliTaq

Gold DNA polymerase (PE Corp., USA), 1x AmpliTaq Gold Buffer, 3mM $MgCl_2$, 5% Dimethylsulphoxide, and 0.2 mM of each dNTP. Thermal-cycling was performed in a 384-well polypropylene plate (ABgene, UK) on a 384 MultiBlock System (Thermo-Hybaid, UK), and entailed an initial 10 minute activation step of 94°C for 10 minutes, followed by 35 cycles of 94°C for 15 seconds and of 55°C for 30s.

Genotyping Assay

To create melt curves we employed the Dynamic Allele-Specific Hybridization (DASH) genotyping method (Genome Research, 2003, 13:916-924). For this, a streptavidin coated nylon membrane was pre-wet in HE buffer (0.05M Hepes, 0.005M EDTA, pH 7.9) and clamped onto an opened PCR plate (post-PCR) enabling the PCR products to be centrifugally transferred onto the membrane (Biotechniques, 2002, 32:1322-1329) to bind and create a macro-array. The membrane was then rinsed in HE buffer and immersed in 0.1M NaOH for 2 minutes to denature the DNA and so remove the non-biotinylated PCR product strand. A further rinse in HE buffer was used to neutralize the pH of the membrane. Probe rs1133104+01P (3'-ROX labeled oligonucleotide 5'-ttctctccCtgtgtgca-3', with the captalized 'C' being the allele-specific base) was then used at 2 pmol/ml in HE buffer to coat the membrane. To drive probe annealing to completion the membrane was placed between glass sheets, heated to 85°C, and allowed to cool to room temperature. The probe was thus annealed to the bound single strand of the PCR product. A final rinse in HE buffer was used to remove excess unbound probe.

To enable assessment of the degree of hybridization between probe and PCR product at any point in time, the membrane was soaked for 3 hours in 40ml HE-buffer containing SYBR Green I

dye at 1:20 000 dilution. This makes it possible to use an induced Fluorescence Resonance Energy Transfer (Genome Research, 2002, 12:1401-1407) interaction between SYBR Green I dye and the ROX label of the probe to generate a fluorescence signal that is related to the existence of double-strand DNA entailing probe and amplified target sequences. To create the melt-curves, the membrane was placed between glass plates under blue light (470nm peak wavelength, suitable for SYBR Green I dye excitation), and its temperature was increased from 30°C to 85°C via a custom-made heating device at a rate of 3°C per minute whilst the fluorescence emitted by each array feature at 630nm (ROX label emission) was measured twice per second via CCD camera imaging. Custom software was then used to quantify the camera images and construct graphic plots of the resulting denaturation melt curves.

Results

The melt-curve graphical plots created for the assayed membrane array features (each feature representing the result for one individual) were a) raw data plots of fluorescence versus temperature (Figure 1), and b) plots of fluorescence versus temperature normalized to an equivalent starting fluorescence value (Figure 2). Additionally, plots were generated to display the negative derivative of fluorescence with respect to temperature versus temperature, and this was done for both the raw and the normalized primary data curves (giving Figure 3 and Figure 4 respectively). In all these plots, the 12 tested DNAs could be seen to fall into one of three distinct clusters; i) curves with a maximal denaturation rate at a relatively low temperature (point A), ii) curves with a maximal denaturation rate at a relatively high temperature (point B), and iii) curves with maximal denaturation rates at both low and high temperatures. Higher

melting temperatures correspond to perfectly matched (more stable) probe-target duplexes, whilst lower melting temperatures correspond to mismatched (less stable) probe-target duplexes, so the three observed curve types equated respectively to i) homozygotes mismatched to the utilized probe sequence (that is, matching the alternate 'A' nucleotide allele), ii) homozygotes matched to the utilized probe sequence ('C' nucleotide allele), and iii) heterozygotes (carrying both the 'C' and the 'A' nucleotide alleles). Grouping lines were drawn manually upon each of the graphical plots (see Figures 1-4) to group the melt curves into the apparent clusters, and these clusters were then assigned to the three genotype categories based upon the interpretation principles described above. Four individuals of each genotype category were thus identified, and these sample assignments were consistent across all of the graphical plots considered. Additional samples may be subsequently examined by the same genotyping method using these established grouping lines for marker rs1133104 to automatically establish the genotype class for any further samples subjected to the same analysis.